

Vision in Natural and Virtual Environments

Mary M. Hayhoe
Center for Visual Science
University of Rochester
Rochester, NY 14627
mary@cvs.rochester.edu

Dana H. Ballard
Computer Science Dept.
University of Rochester
Rochester, NY 14627
dana@cvs.rochester.edu

Jochen Triesch
Cognitive Science Dept.
University of California
La Jolla, CA 92093
triesch@cogsci.ucsd.edu

Hiroyuki Shinoda
Department of Photonics
Ritsumeikan University
Kusatsu Shiga, 525-8577
Japan
hshinoda@se.ritsumei.ac.jp

Pilar Aivar
Psychology Department
University of Oviedo
Oviedo, Asturias 33003
Spain
pdys33@correo.uniovi.es

Brian Sullivan
Center for Visual Science
University of Rochester
Rochester, NY 14627
brians@cvs.rochester.edu

ABSTRACT

Our knowledge of the way that the visual system operates in everyday behavior has, until recently, been very limited. This information is critical not only for understanding visual function, but also for understanding the consequences of various kinds of visual impairment, and for the development of interfaces between human and artificial systems. The development of eye trackers that can be mounted on the head now allows monitoring of gaze without restricting the observer's movements. Observations of natural behavior have demonstrated the highly task-specific and directed nature of fixation patterns, and reveal considerable regularity between observers. Eye, head, and hand coordination also reveals much greater flexibility and task-specificity than previously supposed. Experimental examination of the issues raised by observations of natural behavior requires the development of complex virtual environments that can be manipulated by the experimenter at critical points during task performance. Experiments where we monitored gaze in a simulated driving environment demonstrate that visibility of task relevant information depends critically on active search initiated by the observer according to an internally generated schedule, and this schedule depends on learnt regularities in the environment. In another virtual environment where observers copied toy models we showed that regularities in the spatial structure are used by observers to control eye movement targeting. Other experiments in a virtual environment with haptic feedback show that even simple visual properties like size are not continuously available or processed automatically by the visual system, but are dynamically acquired and discarded according to the momentary task demands.

Keywords: virtual environments, attention, saccadic targeting.

1. WHY STUDY NATURAL TASKS?

In everyday life, visual operations are embedded in the context of ongoing behavior, and depend critically on this immediate context. We know little, however, about how the visual system functions in natural behavior. One of the motivations for studying natural behavior is that the results have immediate applicability to the real world, with consequent applications in a variety of domains including neuropsychological evaluation, and human-computer interactions. At a more theoretical level, it is necessary for us to define exactly what the visual system actually has to do in normal functioning. It is often assumed, for example, that the function of the visual is to identify and locate objects in a scene, and this has guided much of the research effort in visual perception. While this is a reasonable assumption, it characterizes vision as essentially passive, and ignores the dynamic interplay of visual and motor processes on the time scales of tens to hundreds of milliseconds. Additionally, without direct investigation of natural behavior, it is only possible to speak in generalities. Any detailed understanding requires careful examination of situated behavior. To do this, it is important that the experiments provide the same challenges as normal environments. For example, control of eye and hand movements is much simpler for small, two-dimensional displays with the observer in a fixed position. Consequently, the behaviors might be different. Typical experimental displays contain stationary, high contrast, geometric objects. However, in ordinary life, observers must select necessary information from a complex, unsegmented visual array, in the presence of global image motion that is generated by the observer's own movements.

In standard experimental paradigms, a single visual or motor operation is examined over repeated trials. To investigate natural behavior, however, it is necessary to develop a methodology that examines small segments of behavior spanning a few seconds, rather than examining a small time-slice encapsulated in a single trial, and then replicating this time-slice over a series of trials. Examination of behavioral segments, however, allows us to more easily address issues related to the transition from one operation to the next, and the way that a given operation depends on information acquired in previous operations. For example, what controls the transition from

controlling heading direction while walking, to avoiding obstacles, or to what extent is a reaching movement to pick up an object dependent on prior views of that object? Thus difficult questions about selective attention and short-term memory are situated in natural behavior. This requires observations of active observers. For example, viewing a picture of a scene is likely to be very different from acting within that scene simply because the observer needs different information. Similarly, the observer needs to be involved in ongoing behavior where the initiation and timing of visuo-motor operations is controlled by the observer, not the experimenter.

To develop experimental paradigms that investigate these questions, we need to make the environments as natural as possible while still allowing rigorous experimental control. One way of achieving experimental control is to change the information in the display at critical moments during the task, as a probe to identify the information currently being used by the observer. This is the motivation for constructing virtual environments. However, it is also necessary to examine behavior in real environments in order to validate performance in the virtual environment. Behavior in real environments also generates hypotheses about visuo-motor control that can then be translated into the virtual environment for testing.

2. INSIGHTS FROM EYE MOVEMENTS IN REAL ENVIRONMENTS.

Investigation of vision in the natural world has revealed that the pattern and duration of the fixations are highly specialized for each situation. In driving, Land has shown that drivers reliably fixate the tangent point of the curve to control steering around the curve [Land & Lee, 1994]. In cricket, players exhibit very precise fixation patterns, fixating the bounce point of the ball just ahead of its impact [Land & MacLeod, 2000]. A similar pattern is seen in table tennis [Land & Furneaux, 1997]. Ballard et al [1995] found stereotyped fixation patterns in a block-copying task, where subjects acquired different features of the blocks in different fixations. In tea making [Land et al, 1999] and sandwich making [Hayhoe et al, 1999], observers' fixations are tightly linked, step-by-step, with task performance. The duration of the fixations is also different for different tasks. Epelboim et al [1995] showed different fixation durations for tapping and looking at a set of points on a table. Pelz et al [2000] showed different distributions of fixation durations for different stages of a model-building task. In all these cases the visual stimulus is similar throughout the task, but the observer's goals are different. Observers actively select the specific information required for the momentary cognitive goal.

Selection of just the task specific information from a scene is an efficient strategy. As Ullman [1984] has pointed out, it is not possible to anticipate and compute all the information in a scene that might be needed ahead of time. Some kind of selection is necessary to deal with the computational complexity of representing even simple properties of objects and scenes [Ballard, 1991]. Task specific strategies not only circumscribe the information that needs to be acquired, but also allow the visual system to take advantage of the known context to simplify the computation [Ballard et al, 1997]. Ullman [1984] referred to these task-specific computations as "visual routines" and proposed that they are required for

even simple visual operations, such as perception of spatial relations, that have traditionally been considered as automatic. The distinctive feature of visual routines is that they are *procedures* applied on demand. This is in contrast to the more traditional view that such information is automatically extracted by feature detectors driven "bottom-up" or pre-attentively, by the stimulus.

3. VISUAL ROUTINES

To test the view that most visual processing involves active application of specialized routines, we examined performance in a virtual environment, where subjects pick up a set of blocks and move them from one side of a display to the other. We focused on perception of the size of the blocks. This is a simple visual feature that we expect would be automatically extracted when an observer fixates an object. We show, instead, that size information must be actively extracted by the observer, depending on the particular task demands. In addition, once extracted, the information is not necessarily preserved in short term visual memory, but must be recomputed if it is needed at another point in the task. We demonstrated this by changing the size of the block during the movement. Subtle differences in task demands had a big effect on whether this change was visible to the observer, despite the fact that the observer was attending to the block for the entire duration of the movement.

3.1 Virtual Environment

The experimental setup provided both visual and haptic information. The visual display was delivered via a Virtual Research V8 head mounted display, updated at 60 Hz, with a resolution of 640x480 pixels. The stereo image was generated by a Silicon Graphics Onyx II with four 250 MHz processors and two Infinite Reality 2 graphics boards. Head position was monitored with a Polhemus Fastrack, 6 degree-of-freedom, magnetic tracking device. The latency of image updating in this system is about 50 msec. An ASL series 501 infra-red video eye tracker was integrated into the optics of the helmet and used to monitor position of the left eye. The ASL has an accuracy of about 1 degree and temporal resolution of 60 Hz. The ASL signal was recorded in the data stream and in addition, a 30 Hz video record of the display with eye position superimposed was recorded. An image of the observer's eye provided by the ASL was overlaid on the video scene record containing the location of gaze. Two crosshairs on the eye image indicate the tracker's calculation of center of the pupil and corneal reflection. If either of these signals is lost, the corresponding crosshair disappears. This provides a mechanism for checking the scene video for transient track losses and blinks. The movement of the eye can also be seen in the eye image, providing an additional source of information for identifying fixations and measuring their duration.

Haptic information was provided by two opposed Phantom force feedback devices made by SensAble Technology. The Phantom is a specially made research prototype that has 30 times the workspace of that available with the standard commercially available devices. The Phantom is designed for use as follows. The subject's finger is placed in the thimble attached to the end of the device. The system then tracks the position of the fingertip and provides point-force feedback to the

thimble, which is mounted on a 3 degree-of-freedom gimbal so that the force is reflected to the center of the fingertip. With two Phantoms, virtual objects can be gripped between the two thimbles and manipulated within the workspace (40 x 40 x 40 cm, with resolution of 0.03 mm for the standard Phantom). The high-bandwidth interface (updating at 1000 Hz) allows surface and object properties to be modeled and reflected in real time, so virtual objects can be realistically presented to the subject. The object's shape, texture, compliance, static and dynamic friction, can all be modeled with high fidelity. This allows complex, two-finger operations to be performed. A view of the subject with the HMD and using the Phantoms is shown in Plate 1, together with the scene viewed by the subject. Thumb and index finger position is indicated by small spheres as shown in the figure. The transparent block indicates the movement made by the subject to place the block on the conveyor belt on the right, which transports the block out of the field. The crosses indicate eye position sampled at 30 Hz during the movement of the block from left to right.

3.2 Procedure and Results

Subjects picked up bricks of two different heights and placed them on conveyor belts, shown in the color plate. The blocks were either 6x6x8 cm or 6x6x10 cm. At the viewing distance used in the experiment this subtends about 7.6 or 9.5 degrees of visual angle. We can distinguish three phases of the movement for each block: pickup, movement across the field, and put down. In each block of trials, five blocks must be moved onto the conveyor belts. When these are moved, another set of five blocks appears. This was repeated 20 times for each subject.

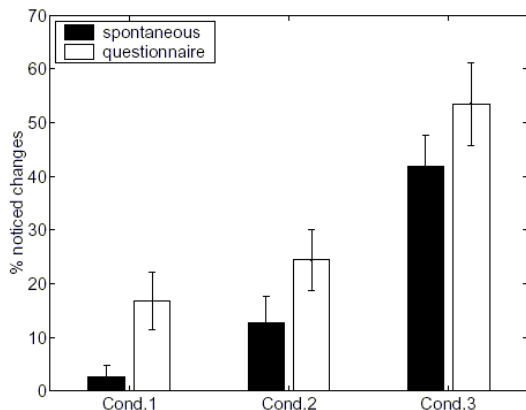


Figure 1. Proportion of changes noticed by subjects.

There were three conditions. In the first, subjects were asked to pick up the blocks in front to back order and place them on the closer conveyor belt. In the second condition, subjects picked up the tall blocks first, then the short ones, placing them on the front belt. In the third condition, subjects picked up the tall blocks and placed them on the front conveyor belt, and then the short ones, putting them on the back belt. In the first condition, there is no particular need to make a size judgment. In the second condition, size must be explicitly judged when the blocks are picked up. In the third condition, size remains relevant to the task at putdown as well. Different

groups of 12 subjects were run in each condition. The goal of the experiment was to see if the condition influenced the visibility of a change in block size during the movement. On 10 percent of the movements, the block changed in height, from small to large or vice versa midway between pickup and putdown. Only the rendered height of the blocks changes, so there are no haptic cues about the change. Subjects were not informed that the blocks might change, but were asked to report anything unusual. If a change was reported, they were asked to report any subsequent changes. In addition, they were questioned at the end of the session. We chose not to inform the subject about possible changes because we wanted to observe the normal distribution of attention during the task, and not induce subjects to do a secondary change detection task, to the extent that this is possible. Figure 1 shows the proportion of reported changes in each of the three conditions. The black bars show the spontaneously reported changes, and the white bars show the reports incremented by reports during the end of session questioning. Although it may be argued that such incidental reporting underestimates the visibility of the changes, the main feature of interest is the greater likelihood of reporting changes in Condition 3, than in either of the other two conditions. The comparison with Condition 2 is particularly interesting. An interpretation of this result is that once the size judgment is made at pickup, it is no longer relevant to the task and the information is no longer represented by the visual system. Subjects need to control grip force and the arm movement, and guide the putdown, and presumably these operations do not require height information.

This effect is not due to a different use of subjects' gaze. In a frame-by-frame analysis of the videotape records of performance, we classified the subject's gaze at the point that the change was made. Frequently subjects make a saccade immediately after pickup to the conveyor belt for guiding the put down movement, and the change occurs during the saccade. The change may also occur just at the onset of the saccade or just as the saccade is terminated. On other trials, however, the subject tracked the block during the movement. Occasionally a blink was occurring at the point of change. The relative frequency of the gaze categories did not differ for the three conditions. In addition, the proportion of changes that were noticed was unaffected by gaze state. Interestingly, this means that a substantial number of changes were unnoticed even when the subject was tracking the block at the moment of the change. This suggests that information about object size is available to the higher levels of the visual system only when it is needed at the current stage of the task. The visual system may not always actively extract or process height information, even though the object is fixated at pick-up and put-down. These first results underscore the dynamic and highly task-specific nature of visual operations. Even very simple visual information such as height seems to require an active visual computation, or visual routine. Thus the visual consequences of a stimulus differ depending on the immediately necessary information.

3.3 Change Blindness and Attention

The current experiment is similar to a number of experiments on the phenomenon of "change blindness", where observers are insensitive to many changes made in scenes either during a saccadic eye movement or some

other masking stimulus. (See [Simons, 2000] for a comprehensive review.) The current experiment provides a unifying principle for understanding what changes are noticed by observers. The inability to notice changes in scenes may have little consequence if that information is not in current use.



in the first place? This problem is particularly challenging in dynamic environments. In the context of driving, for example, it is usually not known ahead of time where the traffic signs are located, or when the car in front might turn. This issue has been described by Ullman [1984] as the “initial access” problem. We refer to it here and

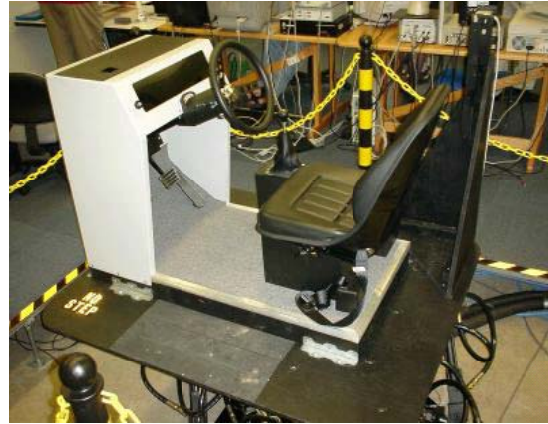


Figure 2: The Performer Town Environment and the Driving Simulator.

Sensitivity to changes has been related to areas rated as having ‘central interest’ in pictures of scenes [O’Regan et al, 1999; Rensink et al, 1997]. However, attentional state in the change blindness experiments is typically not controlled so it is hard to make a definitive link. The current experiments show that sensitivity is directly related to the particular information that observers extract from a scene, and to the precise time when it is extracted. This is consistent with the findings of Henderson & Hollingworth [1999], who showed that subjects are more likely to notice a change in a target made during a saccade to the object than during a saccade away from the object. It also explains O’Regan et al.’s [2000] finding that even objects near fixation can undergo a change without the subjects’ awareness. This might occur if the particular information being extracted differed from that which was changed. O’Regan et al [2000] offer a similar explanation. Simons [2000] points out that there is evidence to support a variety of different accounts of change blindness: subjects might represent information in the initial view before the change, or in the final view, or represent both but fail to compare them. These hypotheses are not exclusive, however, since any one of these possibilities might be in effect, depending on the task demands of the particular experiment.

4. WHAT CONTROLS DEPLOYMENT OF ROUTINES?

The computational advantage of visual routines comes at a cost. Any task driven, or top-down system must deal with the issue of how the particular computations are scheduled. There must be some mechanism for providing the observer with perceptual information that is not on the current agenda. The visual system must balance the selectivity of ongoing task specific computations against the need to remain responsive to novel and unpredictable visual input that may change the task agenda. How do we select what we need to see without knowing what is there

elsewhere [Hayhoe, 2000] as the “scheduling” problem.

Another way to phrase this question is to ask what controls attentional deployment in the normal case of ongoing behavior? Attention can change either endogenously according to the observer’s internal agenda, or exogenously, by stimuli that attract attention. For example, temporal transients are usually thought to attract attention. It is not known what attracts attention in ordinary behavior, and we cannot necessarily rely on bottom up mechanisms to attract attention at the right time. Sudden onsets are rare in ordinary circumstances, and retinal transients are generated primarily by the observer’s own movements. The retinal image is extensive and complex, and observers are typically familiar with the structure of the environment. Attentional distribution will depend on the current behavioral context as well as the properties of the visual stimulus, and this in turn is constantly evolving and not embedded in a set of repeated trials with fixed task demands.

We investigated this problem in the context of a driving task [Shinoda et al, 2001]. No Parking signs were turned into Stop signs for brief periods, and we observed whether this attracted a fixation. We expected that, if subjects needed to actively search for a sign, it should frequently be missed if presented for only a brief interval. If the stimulus itself attracted attention, however, it should be detected whenever it was presented. One way that observers might handle environmental uncertainty in a top-down system is by using learnt knowledge of the probabilistic structure of the environment to initiate task specific computations at likely points [Chun & Yiang 1998, 1999]. Therefore we varied the a priori likelihood of a sign by placing it at either an intersection or in the middle of a block. In addition to this we manipulated the subject’s overall goals. In one condition (F), subjects were simply instructed to follow the car in front of them. In the other condition (F+S) subjects were told in addition to obey the normal traffic rules. If observers rely

on bottom-up scene analysis to initiate a particular visual computation, manipulation of both the a priori likelihood and the subject's goals should have little effect.

4.1 The Driving Simulator

The driving simulator consists of a steering station (originally designed for instructional purposes) mounted on top of a 6 degree-of-freedom hydraulic platform. Steering, accelerator, and brake are instrumented with low-noise potentiometers that modulate signals read by a special-purpose 12-bit A/D board on the SGI's VME bus. Subjects drive in "PerformerTown," an extensible environment designed by Silicon Graphics. The simulator is shown in Figure 2, together with a view of the Performer Town environment. We have added cars and trucks to the environment that move along pre-defined or interactively

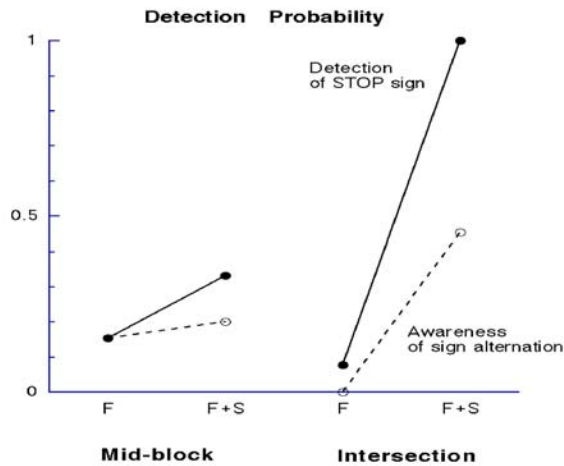


Figure 3. Probability of detecting the stop sign.

controlled paths. These objects can be placed in the environment through a configuration file that makes it easy to customize the environment. The visual environment is also easily modified to simulate lighting at any time of day or night, and the addition of fog. Observers view the display using the V8 helmet as described above, and eye position is monitored using the ASL tracker. Both the video record and the data stream are recorded. Subjects drove along a pre-specified path, following a lead car. In a particular block, a No Parking sign changed into a Stop sign for about 1 sec as the observer approached it. The retinal transient caused by the change was masked by 100 msec blank screens before and after the change.

The results shown indicate that the likelihood of fixating the sign at any point while it is there is heavily modulated by the task and by the location of the sign. The data are plotted in Figure 3. The open symbols show the probability of noticing the change and reporting it at the end of the experiment. The solid symbols show the probability of either fixating the sign, braking, or noticing and reporting. This probability is almost identical to what it would be if fixating alone were used. Thus fixation is the most sensitive measure of a reaction to the manipulation. With the "follow" instruction, observers rarely fixate the sign in either location. When driving normally they invariably respond when it is at an intersection, but miss it 2/3 if the time when it is in the

middle of the block. This suggests that observers must *actively initiate* a search procedure in order to see particular stimuli in the scene. Markedly different fixation patterns were displayed in the two instructions. In the "drive normally" (F+S) condition, Subjects spent much more time fixating in the general region of the intersection than when simply asked to follow the car in front (F condition). This strongly suggests fixation patterns and attentional control in normal vision are learnt.

Our results indicated that observers manage many of the demands of scheduling behaviors in a normal driving environment on the basis of top down mechanisms. That is, driving involves the dynamic application of specialized routines initiated by the observer, whose current goals specify the information that is extracted from a scene, and the time it is extracted. Bottom-up responses appear to be limited in this context. Instead, visibility of traffic signs depends on active search according to an internally generated schedule, and this schedule depends both on the observer's goals and on learnt probabilities about the environment. The experiment suggests that a fuller understanding of the mechanisms of attention, and how attention is distributed in a scene, needs to be situated in the observer's natural behavior.

5. VISUAL MEMORY AND INTEGRATION OF INFORMATION ACROSS FIXATIONS.

If specific information is extracted on each fixation, this suggests that visual operations in different fixations are largely independent. However, there must be some kind of visual memory for information acquired in previous fixations. Thus, we might ask to what extent is visual information acquired in prior fixations needed for normal vision? Observations of eye and hand movements while subjects make sandwiches reveal several aspects of performance that point to the need for some representation of the spatial structure of the scene that is built up over different fixations. Patterns of eye-hand coordination and fixation sequences suggest the need for planning and coordinating movements over a period of a few seconds. For example, subjects frequently look at objects a few seconds before reaching to pick them up or manipulate them. This was observed by Pelz et al [2001] in a hand washing task. In addition, our observations of subjects making peanut butter and jelly sandwiches, reveal that about 30 percent of the reaches are preceded by a fixation on the object within the preceding 5 sec. The relative latency of eye and hand movements for visually guided reaches ranged between plus and minus 1 sec. This broad range of relative latencies suggests that the next eye or hand movement may be planned as much as a second ahead of time, consistent with a visual memory buffer of 1-2 sec for eye-hand coordination. Several fixations may intervene between eye and hand movement to the object, so planning must occur in a representation that is independent of eye position.

Observations of such natural behavior allow only indirect inferences. It is also necessary to test the hypothesis in a situation where the stimulus can be controlled and manipulated. Therefore we have investigated the need for visual memory in targeting movements in another virtual environment.

5.1 Baufix Environment

This environment simulates wooden parts from a toy construction set, called Baufix, shown in Figure 4. The subject's task is to pick up model pieces in a resource region and move them to the other side of the display to make a copy of the model construction at the top of the display. The virtual wooden parts can be picked up and moved using a Polhemus Fastrack 6 degree-of-freedom position sensor as a 3D mouse. Objects are highlighted when the sensor comes in contact with them and the observer can pick the objects up using the space bar of the keyboard. In the experiments described below, the separation between the workspace and model, and resource and model, was about 15 deg. The individual regions subtend about 10 deg. The horizontal field of view of the HMD is 54 deg. As described above, eye and head movements are monitored. In addition, direction of gaze was computed on line and used to change the display contingent on gaze.

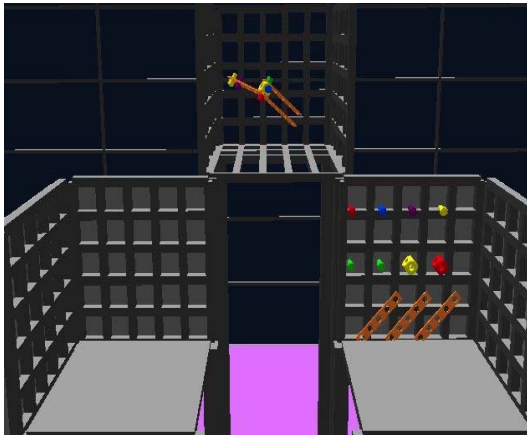


Figure 4. The Baufix Environment. The Model is on the top, the Resource Area is on the right, and the Workspace is on the left.

The focus of this experiment was the saccades to the Resource area for picking up pieces. How are these saccades targeted? Do subjects search for pieces of a particular color, and use this to target the saccade, or do subjects use visual memory for the piece's location? We tested this by scrambling the locations of the pieces when subjects looked away from the Resource area in order to place pieces in the Workspace area. If remembered location is used in targeting the return saccades, such a manipulation should interfere with the process of locating the next piece. Subjects copied the model 10 times. On the first five copies, there were no manipulations of the display. In the next five trials, the Resource pieces were randomly rearranged when the subject's gaze was directed away from the resource following a pickup.

Saccades to the Resource area were categorized using a frame-by-frame analysis of the video records. On many of the pickups (30%), subjects landed in the resource area after a large saccade from the Workspace, and then made several fixations before picking up a piece. In this case it is not clear what the basis of the targeting process is. On some pickups, however, the subject landed first on the old location of a particular piece, that is, its location before

the subject looked away and the pieces were shuffled. The subject then fixated the new location of the piece, and picked it up. We called these Old-New Location pickups. On these trials it seems likely that the initial fixation was made on the basis of spatial memory information. This happened on 21% of the pickups, so it is by no means an exclusive strategy. On about 13% of pickups the saccade landed on the piece that was picked up, without any prior fixations. These were called Direct Fixation pickups. On most of these trials the piece had not moved from its location on the previous visit to the model, so this, too indicates spatial memory use. On a small fraction of these saccades (3%), the piece was in fact in a new location, even though the saccade was made directly to the piece. We can use this fraction as an estimate of the likelihood that the Old-New Location and Direct fixations are chance occurrences. Since the fraction is very small, it seems likely that the Old-New fixations, and most of the Direct fixations (over 30 % of the saccades) are not chance occurrences and were indeed targeted on the basis of spatial information acquired in prior fixations.

These results point to the need for some representation of the spatial structure of the scene that is built up over different fixations. This is in contrast to the work on change blindness, which suggests that very little information is preserved across fixations. O'Regan, Irwin, and coworkers [Irwin, 1991; O'Regan & Levy-Schoen, 1983] suggest that memory across fixations is limited to a coarse description of the identity and location of objects in the scene. The current results suggest that in addition it is necessary to preserve information about spatial structure that is precise enough to serve as a basis for planning and targeting eye and hand movements. Thus the change blindness results may underestimate the extent of visual representations. It would also mean that in normal scenes we may have quite extensive learned representations at least of the invariant aspects of scenes.

6. SUMMARY

In summary, we have developed several virtual environments together with the ability to track observer eye, head, and hand movements. This has allowed us to examine the dynamic microstructure of natural behavior. The predominant finding is that restricted visual information is selectively extracted from the environment for the momentary task and is not preserved in the mental representation unless it is needed. Since this is true for even simple feature information, such as size, the extent of pre-attentive visual processing is probably quite limited. Observers may depend on learnt regularities in the environment to schedule visual search so that unpredictable information is not missed. Despite the limited and transient nature of the representation of the scene, some aspect of the spatial structure of scenes is probably retained in visual memory, because that information is needed for guiding eye and hand movements.

7. ACKNOWLEDGMENTS

This work was supported by NIH grants RR09282 and EY05729.

8. REFERENCES

- BALLARD, D.H. 1991 Animate vision: An evolutionary step in computational vision. *J. of the Inst. of Electronics, Information, and Communication Engineers*, 74, 343-348.
- BALLARD, D, HAYHOE, M. AND PELZ J. 1995 Memory Representations in natural tasks. *Cognitive Neuroscience* 7, 66-80.
- BALLARD, D, HAYHOE, M., POOK, P. AND RAO, R. 1997 Deictic Codes for the Embodiment of Cognition. *Behavioral & Brain Sciences* 20, 723-767.
- CHUN, M. & YIANG, Y. 1998 Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36, 28-71.
- CHUN, M. & YIANG, Y. 1999 Top-down attentional guidance based on implicit learning of visual covariation *Psychological Science*, 10, 360-365.
- EPELBOIM, J., STEINMAN, R., KOWLER, E., EDWARDS, M., PIZLO, Z., ERKELENS, C., AND COLLEWIJN, H. 1995 The function of visual search and memory in sequential looking tasks. *Vision Res.* 35, 3401-22.
- HAYHOE, M. M. 2000 Vision using routines: a functional account of vision. *Visual Cognition*, 7, 43-64.
- HAYHOE, M., LAND, M., AND SHRIVASTAVA, A. 1999 Coordination of eye and hand movements in a normal environment. *Invest. Ophthalmol & Vis. Sci.* 40, S380.
- HENDERSON, J. AND HOLLINGWORTH, A. 1999 The role of fixation position in detecting scene changes across saccades. *Psychological Science*, 10, 438-443.
- D.E. IRWIN, 1991 Information integration across saccadic eye movements, *Cognitive Psychology*, 23, 420-456.
- LAND MF, AND LEE DN 1994 Where we look when we steer. *Nature (Lond.)* 369: 742-744.
- LAND, M. AND FURNEAUX, S. 1997 The knowledge base of the oculomotor system. *Phil. Trans. R. Lond. Soc. B*, 352, 1231-1239.
- LAND, M.F. AND MCCLEOD, P. 2000 The eye movement strategies of batsmen in cricket. *Invest. Ophthalmol. AND Vis. Sci.*, Suppl. 40, S815.
- LAND MF, MENNIE N, AND RUSTED J 1999 Eye movements and the roles of vision in activities of daily living: making a cup of tea. *Perception* 28, 1311-1328.
- J.K. O'REGAN AND A. LEVY-SCHOEN 1983 Integrating visual information from successive fixations: Does trans-saccadic fusion exist? *Vision Research*, 23, 765-769.
- O'REGAN, J. K., RENSINK, R. A. AND CLARK, J. J. 1999 Change-blindness as a result of 'mudsplashes'. *Nature*, 398, 34.
- O'REGAN, J. K., DEUBEL, H., CLARK, J. AND RENSINK, R. A. 2000 Picture changes during blinks: looking without seeing and seeing without looking. *Visual Cognition*, 7, 191-211
- PELZ, J.B., CANOSA, R., BABCOCK, J., KUCHARCZYK, D., SILVER, A., AND KONNO, D., Portable Eyetracking: A Study of Natural Eye Movements, *Proceedings SPIE, Human Vision and Electronic Imaging*, San Jose, CA: SPIE 2000.
- PELZ, J.B. AND CANOSA, R., 2001 Oculomotor Behavior and Perceptual Strategies in Complex Tasks, *Vision Research*. (in press)
- RENSINK, R. A., O'REGAN, J. K. AND CLARK, J. J. 1997 To see or not to see: the need for attention to perceive changes in scenes. *J. Psychol. Sc.*, 8, 368-373
- SHINODA, H., HAYHOE, M., AND SHRIVASTAVA, A. 2001 Attention in natural environments. *Vision Res.* (in press)
- SIMONS, D. J. 2000 Change blindness and visual memory. *A special issue of Visual Cognition*. Psychology Press, Hove, UK,
- ULLMAN, S. 1984 Visual Routines. *Cognition*, 18, 97-157.

